EXHIBIT B

Plant Molecular Biology **38:** 655–662, 1998.

Short communication

Identification of tomato *Lhc* promoter regions necessary for circadian expression

Birgit Piechulla*, Nicole Merforth and Barbara Rudolph[1]

*Universität Rostock, Biochemie, Fachbereich Biologie, Gertrudenstrasse 11a, 18051 Rostock, Germany (*author for correspondence); [1]present address: Institut für Pflanzenbau und Pflanzenzüchtung, Von-Siebold-Strasse 8, 37077 Göttingen, Germany*

Received 21 April 1997; accepted in revised form 14 April 1998

Key words: circadian rhythm, light harvesting complex proteins, *Lhc* gene expression, promoter deletion analysis

Abstract

Expression of the light-harvesting complex protein genes (*Lhc*) is under the control of a circadian clock. To dissect the molecular regulatory components of the circadian clock a promoter deletion analysis of four tomato *Lhc* genes was performed in transgenic tobacco plants. The important 5′-upstream promoter regions are present at different positions relative to the transcription start site of *Lhc* b1*1, b1*2, *Lhc* a3 and *Lhc* a4. A short sequence of 47 nucleotides is necessary for conferring circadian *Lhc* mRNA oscillations. Sequence alignment of the specified promoter regions revealed a novel motif '**CAA**NNNN**ATC**'. This motif is conserved in 5′-upstream regions of clock controlled *Lhc* genes and overlaps with a sequence relevant in phytochrome mediated gene expression.

Abbreviations: Lhc a/b, genes encoding light-harvesting complex proteins (formerly *cab*, chlorophyll *a/b*-binding proteins); DD, continuous darkness; LD, light/dark conditions

Circadian rhythms have been described in nearly every eukaryotic organism as well as in some prokaryotes [19]. Despite the universal appearance and accumulated knowledge about the phenomenon of such rhythms, the molecular mechanism of the circadian clock is still not understood. Evidence is accumulating that indicates that the transmission of the biological clock occurs by feedback and autoregulation [28, 4]. To shed some light on the basic machinery we started to investigate possible components necessary for the circadian mRNA accumulation of the light-harvesting complex proteins (*Lhc*) of plants.

The light-harvesting complexes of plants are organized as protein-pigment units in the thylakoid membranes of chloroplasts which enhance the probability for light quantum absorption and focused channelling of the energy to the photosynthetic reaction centres [1, 16]. The proteins and pigments of these complexes and their respective genes have been intensively investigated in plants. Besides several important findings it turned out that the synthesis of the proteins and the accumulation of respective mRNAs are under the control of a circadian clock [24, 27]. Interestingly, all 19 members of the tomato *Lhc* gene family exhibit this characteristic expression pattern [18]. Based on this similarity a common control mechanism is expected to function at the level of transcription of each tomato *Lhc* gene, for example *cis*- and *trans*-regulatory elements.

Several distinct *cis*-regulatory motifs have been detected for plant genes [22, 29]. A prominent element is the 'ACGT', the G-box, first described by Giuliano *et al.* [13], which is present in the promoter of plant genes encoding very different proteins [31]. However, the *trans*-acting factors binding to this motif belong to the same bZip type [3, 11, 15]. Another sequence ('CCTTATCAT') has been described and correlated with light-responsive expression of several plant genes [14] as well as the 'GATA' motif (I-box). The latter, described by Castresana *et al.* [8], is present

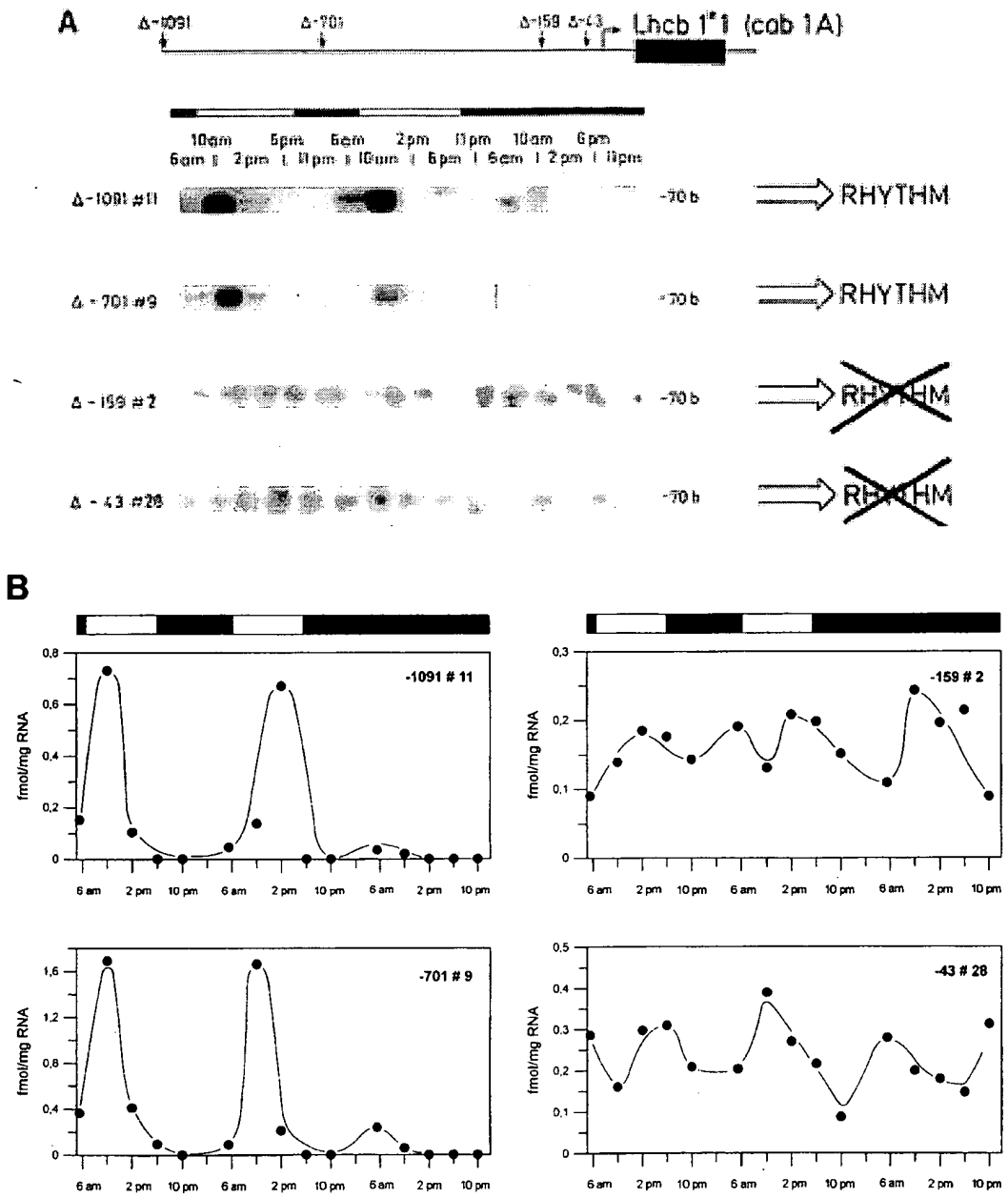

Figure 1. Steady-state mRNA levels of the tomato *Lhc* b1*1 (*cab* 1A) gene in transgenic tobacco. A. Transgenic tobacco plants carrying the deleted tomato *Lhc* gene promoter (Δ-1091, Δ-701, Δ-159, Δ-43) were grown in light/dark (light: 06:00–18:00) and continuous dark conditions. Leaves were harvested at indicated time points and tomato *Lhc* mRNAs were detected with the primer extension technique. The length of the extended primer is 70 bp. B. mRNA levels of deletion constructs were quantitated (fmol per mg RNA) and data of representative tobacco lines are presented.

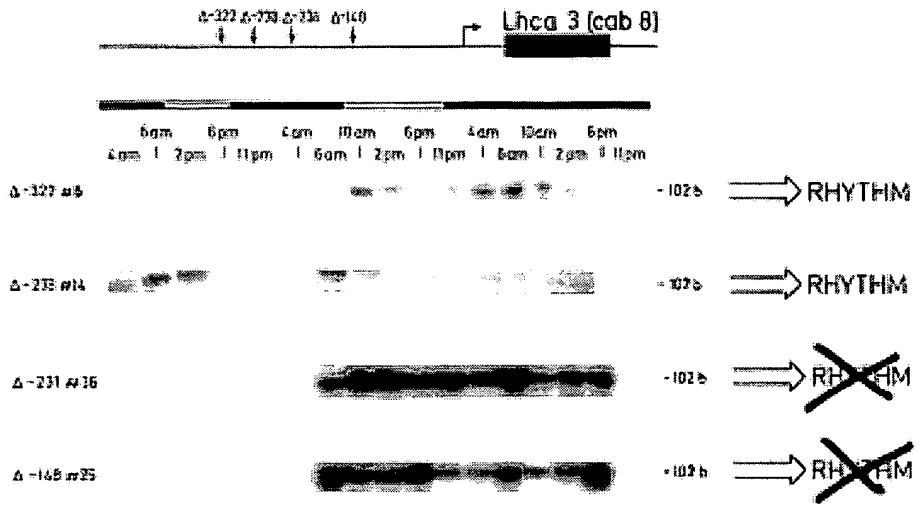

Figure 2. Steady-state mRNA levels of the tomato *Lhc* a3 (*cab* 8) gene in transgenic tobacco. Transgenic tobacco plants carrying the deleted tomato *Lhc* gene promoter (Δ-322, Δ-278, Δ-231, Δ-148) were grown in light/dark (light: 06:00–18:00) and continuous dark conditions. Leaves were harvested at indicated time points and tomato *Lhc* mRNAs were detected with the primer extension technique. The length of the extended primer is 102 bp.

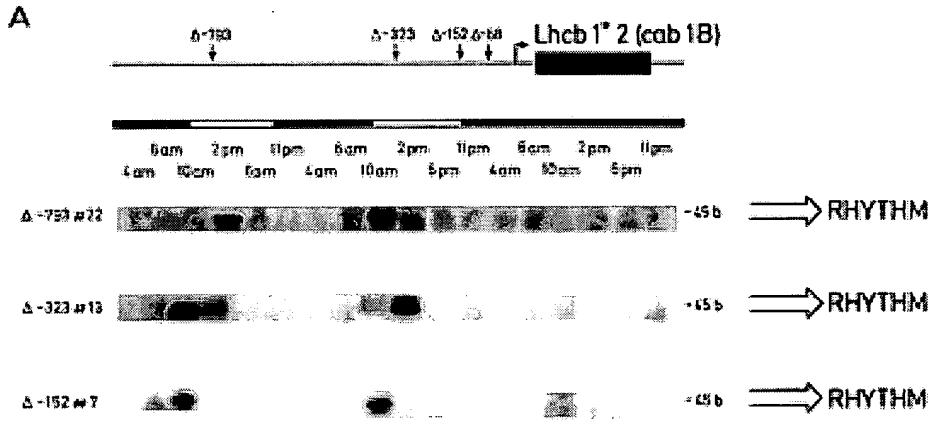

Figure 3. Steady-state mRNA levels of the tomato *Lhc* b1*2 (*cab* 1B) gene in transgenic tobacco. A. Transgenic tobacco plants carrying the deleted tomato *Lhc* gene promoter (Δ-793, Δ-323, Δ-152) were grown in light/dark (light: 06.00–18:00) and continuous dark conditions. Leaves were harvested at indicated time points and tomato *Lhc* mRNAs were detected with the primer extension technique. The length of the extended primer is 45 bp. B. mRNA levels of the deletion constructs were quantitated (fmol per mg RNA) and data of representative tobacco lines are presented.

in the promoter of several light-harvesting complex proteins (LHCP) [12, 25], of the small subunit of Rubisco [9] and in the CaMV 35S promoter [21]. In the case of many *Lhc* gene promoters this motif is two to four times repeated. The nucleotide distances between the 'GATA' sequences are highly conserved [25]. Since the discovery of this motif and its abundant appearance a role as a regulatory unit had been postulated. Variation of the nucleotide sequence of this motif indicated that the 'GATA' element modulates the transcription positively rather than negatively [12]. A correlation of light- and/or tissue-specific expression with the presence of this motif was not observed. A possible function in mediating circadian rhythmicity of *Lhc* mRNA accumulation was discussed since a short promoter fragment of the *Lhc* b1*1 (*cab*2) gene

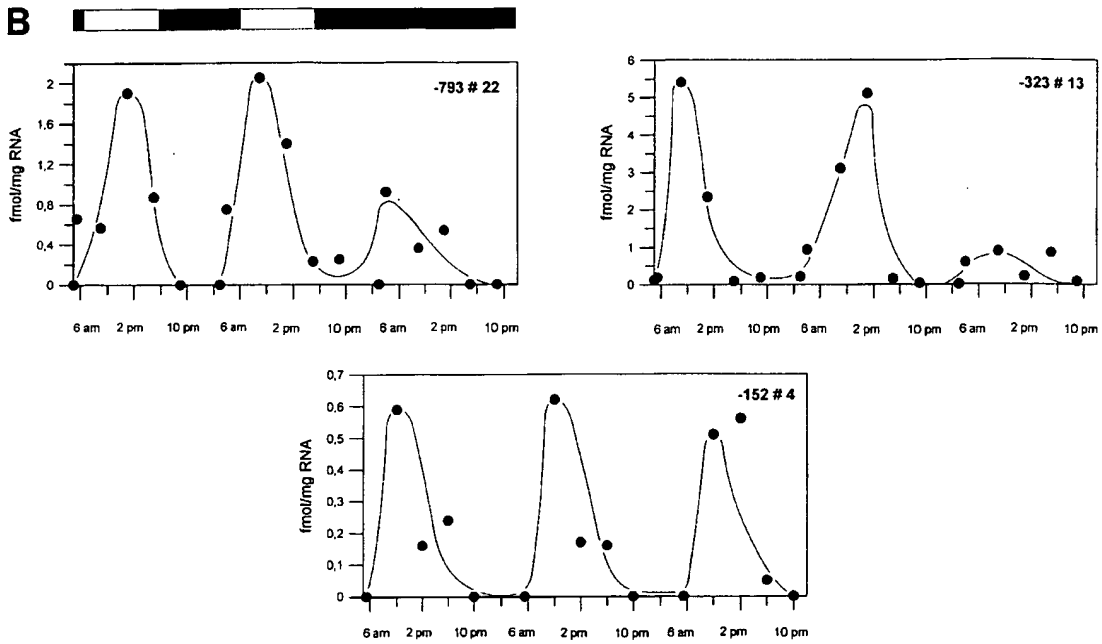

Figure 3. Continued.

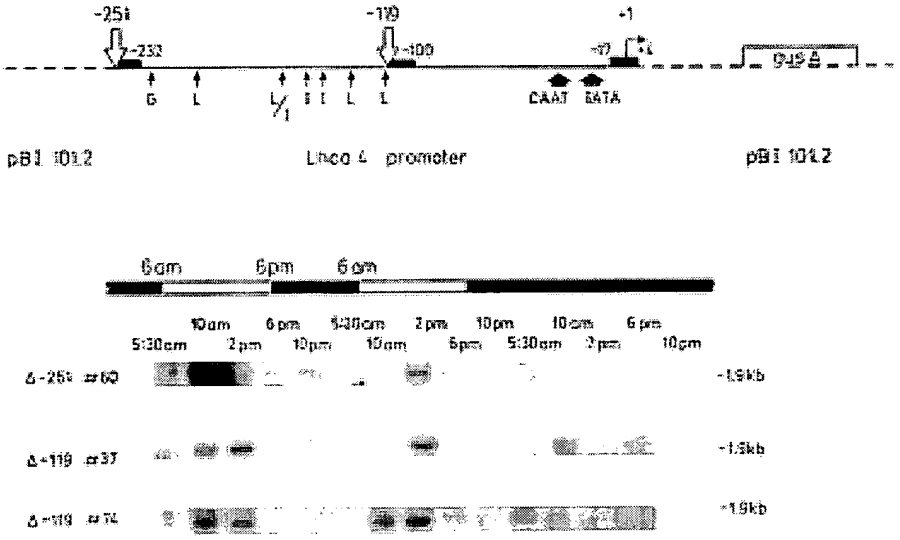

Figure 4. Steady-state mRNA levels of the tomato *Lhc* a4 (*cab* 11) gene in transgenic tobacco. The *Lhc* a4 promoter regions (−251 to +4 or −119 to +4) were cloned into the pBI 101.2 reporter vector from Clontech. Transgenic tobacco plants containing the constructs were grown in light/dark (light: 06:00–18:00) and continuous dark conditions. Leaves were harvested at indicated time points and the glucuronidase mRNA levels were detected by northern blot analysis. The length of the *gus* mRNA is ca. 1.9 kb. For abbreviations, see Figure 5.

of *Arabidopsis thaliana* containing this motif confers circadian *Lhc* expression in transgenic plants [2]. In addition, the sequence 'ACTT' flanking the 'GATA' sequence in tomato was suggested to be a component of the signal transduction chain of the circadian clock [6].

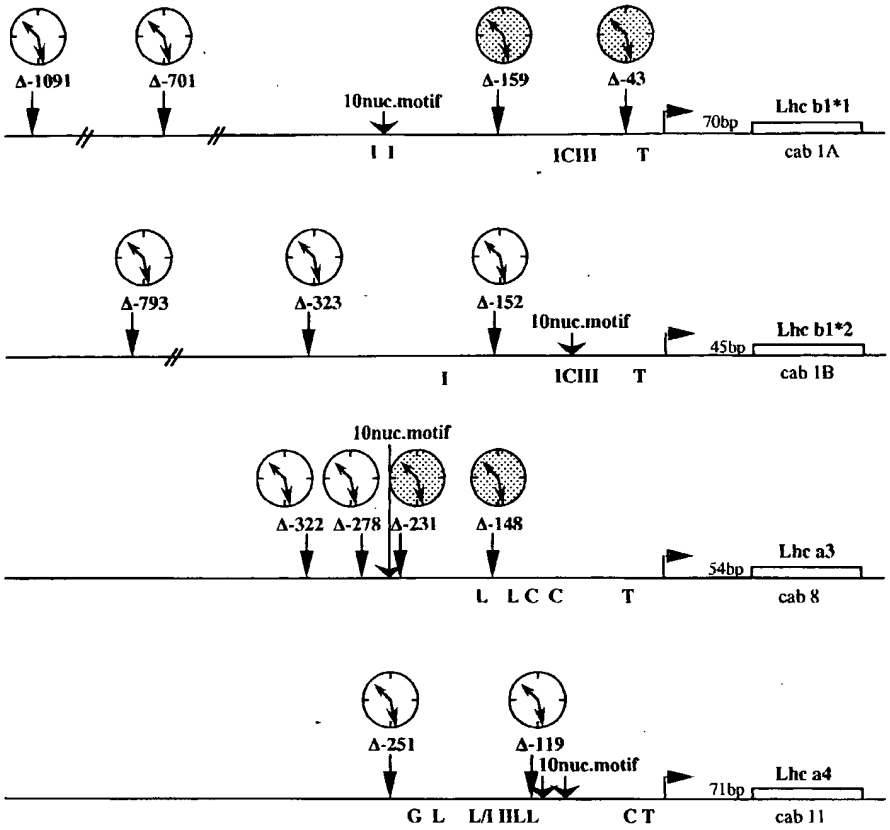

Figure 5. Summary of the deletion analysis. Circadian oscillations of tomato *Lhc* mRNAs in transgenic tobacco are indicated by the 'white' clock. No rhythmic *Lhc* mRNA accumulations were detected in tobacco lines transformed with the tomato *Lhc* b1*1 Δ-159 and Δ-43 and the *Lhc* a3 Δ-231 and Δ-148 deletion constructs and are indicated by the 'shaded' clock. T, 'TATA'; C, 'CCAAT'; I, 'GATA'; L, 'CCTTATCAT'; G,'ACGT'; 10nuc.motif, 'CCANNNNATC'.

Since all members of the tomato *Lhc* gene family express the typical circadian mRNA accumulation pattern a computer-based search was initiated to screen for a conserved *cis*-regulatory element in the 5′-upstream sequences. This analysis was performed with long DNA sequences (e.g. 400 nucleotides) but failed to identify a possible candidate which may be responsible for the mRNA oscillations [25]. Therefore the primary goal of the presented experiments was to explore the 5′-upstream sequences of the tomato *Lhc* genes by promoter deletion analysis in transgenic tobacco plants.

Promoter deletion analysis

Tobacco leaf discs were transformed via *Agrobacterium tumefaciens*-mediated gene transfer with promoter deletion constructs of four tomato *Lhc* genes, *Lhc* b1*1 (*cab* 1A; accession numbers M14445, M30616; promoter: X60922), *Lhc* b1*2 (*cab* 1B, M14443; promoter: X60923), *Lhc* a3 (*cab* 8; X15258), *Lhc* a4 (*cab* 11, X57706). About fifty primary transformants of each deletion construct were regenerated and tested for *Lhc* mRNA expression levels. Individual plant lines with high expression levels were chosen for further analysis.

To find out which of the deleted promoters mediate the circadian oscillations of tomato *Lhc* mRNA, the tobacco transformants were grown in LD and DD conditions. Leaves were harvested at appropriate time points and steady-state mRNA levels were determined by northern blot (Figure 4 for a description of methods, see [24]) or primer extension (Figures 1A, 2, 3A) analysis and quantitated by primer extension analysis (Figure 1B and 3B, for a description of methods, see

Table 1. Sequence motif presence in 5′-upstream regions of *Lhc* genes.

Lycopersicon esculentum	(−701 to −159)	*Lhc* b1*1	−254	CAAAGATATC
	(−152 to −1)	*Lhc* b1*2	−87	CAATGAGATC*
	(−278 to −231)	*Lhc* a3	−237	CAAGAGTATC
	(−119 to −1)	*Lhc* a4	−116	CAACTCAATC
			−92	CAAAAAATC
Triticum aestivum	(−211 to −90)	*Lhc* b1	−176	CAAGAGTATC
			−140	CAATGGCATC
Arabidopsis thaliana	(−111 to −74)	*Lhc* b1*1	−103	CAAAAATC
			−87	CAATGAATGA*
	(−900 to −1)	*Lhc* b1*2	−371	CAATGGAATC
			−253	CAAATAAGTTATC
			−174	CAATGAAAAAATC
			−106	CAAAATC

*Sequence overlaps with the 'CAAT' box.

[18, 25]. The mRNA levels of the tomato *Lhc* b1*1 deletions Δ-1091 (Nos. 4, 11, 38) and Δ-701 (Nos. 4, 9, 26) clearly oscillate in LD with a period of ca. 24 h. Only very little or no tomato *Lhc* mRNA is detectable at 06:00, maximum mRNA levels were reached at 10:00 and levels decrease thereafter (Figure 1A and B). The phase of the circadian rhythms are not altered in different deletion constructs. In constant darkness the amplitudes are significantly reduced but transcripts increase after the night trough at the appropriate time may indicate that an endogenous oscillator influences the expression of the tomato *Lhc* gene.

In contrast to the long promoter constructs it is very unlikely that the tobacco lines with the short promoter regions of *Lhc*b1*1 Δ-159 (Nos. 2, 42, 47) and Δ-43 (Nos. 24, 28, 33, 34) exhibit circadian mRNA accumulation patterns (representative data in Figure 1). The mRNA levels fluctuate only twofold at low expression levels, however a constant period length of ca. 24 h was not measured in LD and DD, particularly well documented after quantitation (Figure 1B, right panels). These results clearly show that the short 5′-upstream regions are sufficient for a basal mRNA accumulation and strongly suggests that sequences upstream of −159 are necessary for circadian mRNA accumulation of the tomato *Lhc* b1*1 gene. Furthermore, it is interesting to note that the expression level of the short constructs do not decrease in DD, while this is usually observed for complete *Lhc* genes which are in their native 3′ and 5′ nucleotide surrounding.

With the *Lhc* a3 deletion constructs similar results were obtained as with *Lhc* b1*1 (Figure 2). Deletion Δ-322 (No. 6) and Δ-278 (No. 14) exhibited circadian *Lhc* mRNA accumulation while the mRNA of the deletion constructs Δ-231 (Nos. 24, 36) and Δ-148 (Nos. 16, 25) reached almost constant levels in LD and DD and no oscillations with a defined period length could be observed. Based on these data it is likely that a region of 47 nucleotides (−278 to −231) is necessary for circadian *Lhc* a3 mRNA oscillations. The low expression levels of the *Lhc* a3 gene in transgenic tobacco prevented a quantitation with the primer extension analysis.

Investigation of the *Lhc* b1*2 (Figure 3) and *Lhc* a4 (Figure 4) genes in tobacco revealed circadian *Lhc* mRNA accumulations for constructs that contain long as well as short 5′-upstream regions (*Lhc* b1*2:Δ-793, Nos. 11, 22; Δ-323, Nos. 13, 17; Δ-152, Nos. 1, 4, 21, 24, *Lhc* a4: Δ-251, No. 60, Δ-119, No. 37, 74). The mRNA constructs of the *Lhc* b1*2 gene could be quantitated with the primer extension analysis and representative results are depicted in Figure 3B. A circadian oscillatory pattern is obvious for each deletion construct. It should be noted that the short 5′-upstream regions of *Lhc* b1*2 and *Lhc* a4 are apparently sufficient to confer circadian rhythmicity. This is a surprise since in contrast such short promoter regions of the *Lhc* b1*1 and *Lhc* a3 genes exhibit an almost constant expression pattern (Figures 1 and 2). It therefore can be concluded that different 5′-upstream regions (relative to the transcription start sites) of the tomato *Lhc* genes are necessary for conferring circadian rhythmicity. The results of all deletion constructs are summarized in an overview in Figure 5.

Sequence comparison

The promoter deletion analysis indicated regions which are necessary for the transmission of the circadian clock in transgenic plants. If a conserved 'clock' regulatory element exists it should be localized within these respective regions. Therefore the following 5′-upstream sequences were aligned: *Lhc* b1*1, −701 to −159; *Lhc* b1*2, −152 to −1; *Lhc* a3: −278 to −231; *Lhc* a4, −119 to −1. A novel sequence motif appears to be conserved in all four tomato *Lhc* 5′-upstream regions (Table 1). The sequence is '**CAA**NNNN**ATC**', it is composed of three conserved nucleotides, four variable nucleotides followed by three conserved nucleotides. Identical or very similar motifs are present in short 5′ promoter regions of circadian controlled *Lhc* genes of *Arabidopsis thaliana* (*Lhc* b1*1 (*cab* 2), −111 to −74 [2, 7]; *Lhc* b1*2 (*cab* 3), −371 [23] and *Triticum aestivum* (*Lhc* b1 (*cab* 1), −211 to −90 [10]). Consistent with the hypothesis that this motif plays a role in transmitting circadian rhythmicity is its absence in the promoter of the non-circadian expressed *Lhc* b gene of *Pinus contorta* (−1000 to +1 [5]). Also interesting in this context are the investigations of Wang *et al.* [30]. Following the phytochrome-regulated light signal transduction pathway of the *A. thaliana Lhc* b1*3 gene a specific protein, CCA1, a myb-related transcription factor, was identified. Interestingly, the binding site of this transcription factor overlaps with the motif which we found to be conserved in the 5′-upstream regions that are necessary for circadian expression of the four tomato *Lhc* genes. These findings support the emerging idea that light perception and clock function are closely related, a hypothesis recently stated by Kay [17]. His conclusion was based on the presence of the PAS domains in clock-related proteins such as PER (period, *Drosophila*) and WC (white collar, *Neurospora*) as well as in photosensory proteins such as phytochrome (PHY, *plants*) and the bacterial blue-light receptor (PYP, photoactive yellow protein [20]). Isolation and characterization of the *trans*-regulatory factor(s) binding to the novel '**CAA**NNNN**ATC**' motif are on the way and will bring further knowledge regarding the hypothesis that the signal transduction pathways of the light and the circadian clock cross-talk or converge.

Acknowledgements

We thank Dr E. Pichersky (Ann Arbor, MI) for providing us with the tomato *Lhc* a3 and *Lhc* a4 genomic clones and unpublished promoter sequences. Thanks are also due to Dr U. Sonnewald and coworkers (Gatersleben, Germany) who performed the plant transformations. We thank Mrs U. Schäfer for the deletion constructs of *Lhc* a3, Mr D. Dreesmann for inital investigations of some transgenic lines and M. Nolte for sequencing the promoter of *Lhc* b1*1 of tomato. B.P. was supported by the DFG (153/8-2).

References

1. Anderson JM, Andersson B: The dynamic photosynthetic membrane and regulation ofsolar energy conversion. Trends Biochem Sci 13: 351–355 (1988).
2. Anderson SL, Teakle GR, Martino-Catt SJ, KaySA: Circadian clock- and phytochrome-regulated transcription is conferred by a 78 bp *cis*-acting domain of the *Arabidopsis cab* 2 promoter. Plant J 6: 457–470 (1994).
3. Armstrong GA, Weisshaar B, Hahlbrock K: Homodimeric and heterodimeric leucine zipper proteins and nuclear factors from parsley recognize diverse promoter elements with 'ACGT' cores. Plant Cell 4: 525–537 (1992).
4. Aronson BD, Johnson KA, Loros JJ, Dunlap JC: Negative feedback defining a circadian clock: Autoregulation of the clock gene frequency. Science 263: 1578–1584 (1994).
5. Barrett JW, Beech RN, Dancik BP, Strobeck C: A genomic clone of a type I *cab* gene encoding a light harvesting chlorophyll *a/b* binding protein of photosystem II identified from lodgepole pine. Genome 37: 166–172 (1994).
6. Borello U, Ceccarelli E, Giuliano G: Constitutive, light-responsive, and circadian clock-responsive factors compete for the different I-box elements in plant-regulated promoters. Plant J 4: 611–619 (1993).
7. Carre IA, Kay SA: Multiple DNA-protein complexes at a circadian-regulated promoter element. Plant Cell 7: 2039–2051 (1995).
8. Castresana C, Staneloni R, Malik VS, Cashmore AR: Molecular characterization of two clusters of genes encoding the type I *cab* polypeptides of PSII in *Nicotiana plumbaginifolia*. Plant Mol Biol 10: 117–126 (1987).
9. Donald RGK, Schindler U, Batschauer A, Cashmore AR: The plant G-box promoter sequence activates transcription in *Saccharomyces cerevisiae* and is bound in vitro by a yeast activity similar to GBF, the plant G-box binding factor. EMBO J 9: 1727–1735 (1990).
10. Fejes E, Pay A, Kanevsky I, Szell m, Adam E, Kay SA, Nagy F: A 268 bp upstream sequence mediates the circadian clock-regulated transcription of the wheat *cab* 1 gene in transgenic plants. Plant Mol Biol 15: 921–932 (1990).
11. Foster R, Izawa T, Chua NH: Plant bZIP proteins gather at 'ACGT' elements. FASEB J 8: 192–199 (1994).
12. Gidoni D, Brosio P, Bond-Nutter D, Bedbrook J, Dunsmuir P: Novel *cis*-acting elements in *Petunia cab* gene promoters. Mol Gen Genet 215: 337–344 (1994).

13. Giuliano G, Pichersky E, Malik VS, Timko MP, Scolnik PA, Cashmore AR: An evolutionary conserved protein binding sequence upstream of a plant light-regulated gene. Proc Natl Acad Sci USA 85: 7089–7093 (1988).
14. Grob U, Stüber K: Discrimination of phytochrome-dependent, light-inducible from non-light-inducible plant genes. Prediction of a common light-responsive element (LRE) in phytochrome-dependent, light-inducible plant genes. Nucl Acids Res 15: 9957–9973 (1987).
15. Izawa T, Foster R, Chua NH: Plant bZIP protein binding specificity. J Mol Biol 230: 1131–1144 (1993).
16. Jansson S: The light-harvesting chlorophyll *a/b* binding proteins. Biochim Biophys Acta 1184: 1–19 (1994).
17. Kay S: PAS, present, and future: clues to the origins of circadian clocks. Science 276: 753–754 (1997).
18. Kellmann JW, Merforth N, Wiese M, Pichersky E, Piechulla B: Concerted circadian oscillations in transcript levels of nineteen *Lha/b* (*cab*) genes in *Lycopersicon esculentum* (tomato). Mol Gen Genet 237: 439–448 (1993).
19. Kondo T, Strayer CA, Kulkarni RD, Taylor W, Ishiura M, Golden SS, Johnson CH: Circadian rhythms in prokaryotes: luciferase as a reporter of circadian gene expression in cyanobacteria. Proc Natl Acad Sci USA 90: 5672–5676 (1993).
20. Lagarias D.M., Shu-Hsing W., Lagarias J.C: A typical phytochrome gene structure in the green alga *Mesotaenium caldariorum*. Plant Mol Biol 29: 1127–1142 (1995).
21. Lam E, Chua NH: ASF-2: A factor that binds to the cauliflower mosaic virus 35S promoter and a conserved GATA motif in *cab* promoters. Plant Cell 1: 1147–1156 (1989).
22. Manzara T, Carrasco P, Gruissem W: Develpmental and organ-specific changes in promoter DNA-protein interactions in the tomato *rbc* S gene family. Plant Cell 3: 1305–1316 (1991).
23. Millar A, Kay SA: Circadian control of *cab* gene transcription and mRNA accumulation in *Arabidopsis*. Plant Cell 3: 541–550 (1991).
24. Piechulla B, Gruissem W: Diurnal mRNA fluctuations of nuclear and plastid genes in developing tomato fruits. EMBO J 6: 3593–3599 (1987).
25. Piechulla B, Kellmann JW, Pichersky E, Schwartz E, Förster HH: Determination of steady-state mRNA levels of individual chlorophyll *a/b* binding protein genes of the tomato *cab* gene family. Mol Gen Genet 230: 413–422 (1991).
26. Piechulla B: 'Circadian clock' directs the expression of plant genes. Plant Mol Biol 22: 533–542 (1993).
27. Riesselmann S, Piechulla B: Diurnal and circadian light-harvesting complex and quinone B-binding protein synthesis in leaves of tomato (*Lycopersicon esculentum*). Plant Physiol 100: 1840–1845 (1992).
28. Sassone-Corsi P: Rhythmic transcription and regulatory loops: winding up the biological clock. Cell 78: 361–364 (1994).
29. Schindler U, Cashmore AR: Photoregulated gene expression may involve ubiquitous DNA binding proteins. EMBO J 9: 3415–3427 (1990).
30. Wang Z-Y, Kenigsbuch D, Sun L, Harel E, Ong M.S, Tobin E.M: A myb-related transcription factor is involved in the phytochrome regulation of an *Arabidopsis Lhc* b gene. Plant Cell 9: 491–507 (1997).
31. Williams ME, Foster R, Chua NH: Sequences flanking the hexameric G-box core 'ACGTTG' affect the specificity of protein binding. Plant Cell 4: 485–496 (1992).